On the maximum entropy principle in non-extensive thermostatistics

Jan Naudts

Departement Fysica, Universiteit Antwerpen,

Universiteitsplein 1, 2610 Antwerpen, Belgium

E-mail: Jan.Naudts@ua.ac.be.

May 2004

Abstract

It is possible to derive the maximum entropy principle from thermodynamic stability requirements. Using as a starting point the equilibrium probability distribution, currently used in non-extensive thermostatistics, it turns out that the relevant entropy function is Rényi's α-entropy, and not Tsallis' entropy.

1 Introduction

The maximum-entropy principle has played an essential role in the development of non-extensive thermostatistics, starting with Tsallis' paper [1] in 1988. The Tsallis formalism has known several major revisions without reaching a stable and satisfactory status. Origin of the difficulties might be that the maximum-entropy principle is used as a postulate of statistical mechanics, while it is rather a result that can be derived from first principles. Below is used that the maximum-entropy principle is a consequence of the statement that free energy is minimal at equilibrium. The definition of what equilibrium is, what free energy is, comes first. The maximum-entropy principle follows. This point of view differs from that of Tsallis literature where equilibrium is by definition the state maximizing some entropy functional under certain constraints.

This change of viewpoint has some important consequences. Given a family of temperature dependent probability distributions one can calculate the entropy functional which is needed to reproduce the given probability

distributions. In the present paper the probability distributions used in recent Tsalllis literature are taken as starting point. Surprisingly, the resulting entropy functional is *not* Tsallis entropy but is the α-entropy of Rényi. This result agrees with the known fact [2] that the maximum entropy principle in combination with Rényi's entropy reproduces the equilibrium probability distributions of non-extensive thermostatistics. Note that there exists an extensive literature on Rényi's α-entropies. Some of the references in the context of non-extensive thermostatistic are [2, 3, 4, 5].

The next section serves to fix notations and starts with the probability distributions that are found as equilibrium states in recent Tsallis literature. Section 3 introduces thermodynamic entropy $S(U)$ as a function of average energy U. It is obtained from average energy by integrating the well-known thermodynamic definition of temperature. In Section 4 the requirement of thermodynamic stability is worked out. In Section 5 the maximum entropy principle is derived. The final section contains a short discussion of results. Some of the calculations are given in Appendix.

2 Equilibrium probability distributions

For sake of simplicity notations refer to a classical gas of n particles enclosed in a container of volume V in ν-dimensional space. The energy functional is denoted $H(x)$, with x a point of phase space Γ.

Average energy

$$U(\beta) = \int_\Gamma \mathrm{d}x\, p_\beta(x) H(x) \tag{1}$$

is calculated using the distribution [6, 7]

$$p_\beta(x) = \frac{1}{Z(\beta)} \left[1 - (1-q)\beta[H(x) - U(\beta)]\right]_+^{q/(1-q)}. \tag{2}$$

In this expression Z is the normalization

$$Z(\beta) = \int_\Gamma \mathrm{d}x\, \left[1 - (1-q)\beta[H(x) - U(\beta)]\right]_+^{q/(1-q)}. \tag{3}$$

The notation $[u]_+ = \max\{0, u\}$ is used. Equation (2) can be written as

$$p_\beta(x) = \frac{1}{Z^*(\beta^*)} \left[1 - (1-q)\beta^* H(x)\right]_+^{q/(1-q)} \tag{4}$$

with

$$\beta^* = \frac{\beta}{1 + (1-q)\beta U(\beta)}. \tag{5}$$

In the latter form normalization is straightforward and average energy can be calculated

$$U(\beta) = \frac{1}{Z^*(\beta^*)} \int_\Gamma dx \; [1 - (1-q)\beta^* H(x)]]_+^{q/(1-q)} \, H(x) \tag{6}$$

The pair of equations (5, 6) should have a unique solution $U(\beta)$. In addition, this solution should be a decreasing function of β. See Appendix A for a discussion of the existence of solutions.

3 Thermodynamic entropy

The basic thermodynamic quantity is entropy $S(U)$ as a function of energy U. It is used to define inverse temperature β by

$$\beta = \frac{dS}{dU}. \tag{7}$$

In the present context $U(\beta)$ is given. Hence $S(U)$ can be obtained by a simple integration of (7). This is done in Appendix B. The result is

$$S(U(\beta)) = \ln\left[\left(\frac{m}{h}\right)^{\nu n} Z(\beta)\right]. \tag{8}$$

Here, m is the mass of each particle, and h is a constant inserted for dimensional reasons (units are used in which $k_B = 1$). The integration constant has been chosen so that for the ideal gas in the limit $q = 1$ one obtains the standard result

$$S(U) = \frac{\nu n}{2} \ln \frac{2\pi em}{h^2} U V^{2/\nu} - \frac{\nu n}{2} \ln \frac{\nu n}{2}. \tag{9}$$

4 Thermodynamic stability

Thermodynamic stability of equilibrium states requires [8] in the first place that

$$\frac{\partial^2 S}{\partial U^2} \leq 0. \tag{10}$$

This condition boils down to the requirement that $U(\beta)$ is a decreasing function of β. This is assumed to be the case, but cannot be proved under general conditions for the family of probability distributions (2).

If (10) is satisfied for all U then $S(U)$ can be replaced by its Legendre transform, which is called Helmholtz free energy F, and which historically is defined by

$$F = U - TS(U) \tag{11}$$

with temperature T given by $T = 1/\beta$. Free energy F is a concave function of temperature T. Its first derivative satisfies

$$\frac{\mathrm{d}F}{\mathrm{d}T} = -S. \tag{12}$$

Thermodynamic stability is more than just (10). The system should be stable against all possible perturbations. For that reason the free energy should be concave, not only as a function of temperature but in a much larger parameter space. To formulate this requirement in a consistent way it is easier to work with $\beta F = \beta U - S$ instead of with F. Note that

$$\frac{\mathrm{d}}{\mathrm{d}\beta}\beta F = U. \tag{13}$$

Hence the condition of concavity of $F(T)$ is equivalent with concavity of βF as a function of β. The quantity βF depends on both β and on the choice of energy function $H(x)$, but only the combination $\beta H(x)$ occurs. Hence the requirement that βF is concave in β can be extended to the requirement that βF is a concave function over the space of all possible $\beta H(x)$. The latter requirement is too strong, because of its global nature, and should be weakened by requiring local concavity of βF. A tangent plane to βF in the point βH is given by the linear map

$$\beta' H' \rightarrow \beta F + \beta'\langle H'\rangle - \beta\langle H\rangle \tag{14}$$

where averages are equilibrium averages w.r.t. β, H. The requirement for stability of the system with energy functional $H(x)$ at inverse temperature β is therefore that

$$\beta' F' \leq \beta F + \beta'\langle H'\rangle - \beta\langle H\rangle \tag{15}$$

for all possible choices of β', H'. Note that (15) can be written as

$$S' - \beta'U' \geq S - \beta\langle H'\rangle. \tag{16}$$

In this expression, the averages $\langle \cdot \rangle$ are still calculated w.r.t. the equilibrium probability distribution of the system with energy functional $H(x)$ at inverse temperature β.

5 Maximum entropy principle

The origin of the maximum entropy principle lies in the requirement of thermodynamic stability. Indeed, (16) can be read as the statement that the quantity $S - \beta \langle H' \rangle$ is maximal when $\beta = \beta'$ and $H = H'$, i.e., when for the calculation of free energy the correct probability distribution is used. This is the maximum entropy principle. In the canonical formalism of statistical physics, based on the Boltzmann-Gibbs distribution, systems with a finite number of degrees of freedom are generically stable. Hence it is acceptable to interchange primed and unprimed systems. In the more general context of non-extensive thermostatistics this interchange is not acceptable, since it is known (see e.g. [9]) that instabilities can occur. One concludes that the maximum entropy principle holds only under the condition that the variation of $S - \beta \langle H' \rangle$ is restricted to those pairs of inverse temperature β and energy functional $H(x)$ which are thermodynamically stable. Of course, this statement is of an academic rather than practical value. But it shows that maximum entropy should not be introduced in an axiomatic manner. It also gives somewhat more freedom in deriving a variational principle which is useful for practical purposes because it is not necessary to prove that the variational solution corresponds with a maximum. Indeed, in principle the variation should be restricted to the equilibrium distributions of stable systems, and by definition of stability these will have a lower free energy than the variational solution.

Let us now introduce an entropy functional which reproduces (2) by means of the maximum entropy principle. Inspection of (8) learns that the correct *ansatz* for entropy as a function of probability distribution is

$$I(p) = -\ln \left(\frac{m}{h} \right)^{\nu n} \left[\int_\Gamma \mathrm{d}x \, p(x)^{1/q} \right]^{q/(1-q)} \tag{17}$$

This expression is proportional to Rényi's α-entropy with $\alpha = 1/q$. One verifies immediately, using the expressions from Section 2, that $I(p_\beta) = S(U(\beta))$. Variation of

$$\mathcal{L}(p) = I(p) - \beta \int_\Gamma \mathrm{d}x \, p(x)[H(x) - U(\beta)] - \alpha \left[\int_\Gamma \mathrm{d}x \, p(x) - 1 \right] \tag{18}$$

gives

$$p(x) = \frac{1}{Z(\beta)} \left[-(1-q)\alpha - (1-q)\beta[H(x) - U(\beta)] \right]_+^{q/(1-q)} \tag{19}$$

with

$$Z(\beta) = \left[\int_\Gamma \mathrm{d}x \, p(x)^{1/q} \right]^{-q/(1-q)} \tag{20}$$

5

Integrating $p(x)^{1/q}$, as given by (19), gives

$$
\begin{aligned}
Z(\beta) &= \int_{\Gamma} \mathrm{d}x \ [-(1-q)\alpha - (1-q)\beta[H(x) - U(\beta)]]_{+}^{1/(1-q)} \\
&= -(1-q)\alpha \int_{\Gamma} \mathrm{d}x \ [-(1-q)\alpha - (1-q)\beta[H(x) - U(\beta)]]_{+}^{q/(1-q)} .
\end{aligned}
\tag{21}
$$

Hence $\alpha = -1/(1-q)$ results in a correctly normalized probability distribution. This shows that (19) coincides with (2). One concludes that there exists a maximum entropy principle which produces (2) as the equilibrium probability distributions and is such that the attained maximum value coincides with thermodynamic entropy.

6   Discussion

This paper contains one message and one result. The message is that the maximum entropy principle is a consequence of the requirement of thermodynamic stability. In the context of the Boltzmann-Gibbs distribution this is an evidence. In that case systems with a finite number of degrees of freedom are automatically stable. This allowed Jaynes [10] to attribute an axiomatic status to the maximum entropy principle. In the current version of Tsallis thermostatistics systems with a finite number of degrees of freedom are *not* automatically stable. The conclusion of the present paper is therefore that maximum entropy should not be used in an axiomatic manner, but should be derived from the requirement of thermodynamic stability.

The result of the present paper is that the equilibrium probability distributions, as found in recent Tsallis literature, can be derived by means of the maximum entropy principle. However, the relevant entropy functional is Rényi's α-entropy rather than Tsallis' entropy functional. Of course, there is an easy connection between the two entropies [11]. But the present paper shows that a thermodynamically correct definition of temperature leaves no choice but to use Rényi's entropy instead of that of Tsallis.

Mainstream non-extensive thermostatistics is based on a variational principle which involves escort probability distributions [6, 7]. The main objection, from the present point of view, against these papers is that they do not reproduce thermodynamic entropy (8). As a consequence, the mainstream derivation of (2) has problems with a correct definition of temperature. Bashkirov, in the conclusions of his paper [4], hopes that one "can abandon the use of escort probabilities for calculations of average values in favor of original distributions." This statement seems to be exaggerated. We

know that escort probabilities are important in fractal physics [12]. As shown in [13, 14, 15] they also play an important role in generalized thermostatistics. It only happens that these escort probabilities are not involved in the formulation of the maximum entropy principle.

Appendix A

In this appendix the solution of (5, 6) is discussed. Elimination of $U(\beta)$ gives

$$
\begin{aligned}
\frac{1}{\beta^*} &= \frac{1}{\beta} + (1-q)U(\beta) \\
&= \frac{1}{\beta} + \frac{1}{\beta^*} \frac{1}{Z^*(\beta)} \int_{\Gamma} dx \ [1 - (1-q)\beta^* H(x)]]_+^{q/(1-q)} (1-q)\beta^* H(x) \\
&= \frac{1}{\beta} + \frac{1}{\beta^*} - \frac{1}{\beta^*} \frac{1}{Z^*(\beta)} \int_{\Gamma} dx \ [1 - (1-q)\beta^* H(x)]]_+^{1/(1-q)} .
\end{aligned}
\tag{A1}
$$

This simplifies to

$$
\beta^* = \beta \frac{\int_{\Gamma} dx \ [1 - (1-q)\beta^* H(x)]]_+^{1/(1-q)}}{\int_{\Gamma} dx \ [1 - (1-q)\beta^* H(x)]]_+^{q/(1-q)}} .
\tag{A2}
$$

See [16] for methods to prove the existence and uniqueness of the solution β^* of this equation. Then average energy is given by

$$
U(\beta) = \frac{1}{1-q} \left(\frac{1}{\beta^*} - \frac{1}{\beta} \right) .
\tag{A3}
$$

Thermodynamic stability requires that $U(\beta)$ is a decreasing function of β. There is no general argument why this should always be the case.

Appendix B

In this appendix entropy $S(U)$ is calculated. One has

$$
\begin{aligned}
0 &= (1-q)\beta \left[\langle H \rangle - U(\beta) \right] \\
&= \frac{1}{Z(\beta)} \int_{\Gamma} dx \ [1 - (1-q)\beta[H(x) - U(\beta)]]_+^{q/(1-q)} (1-q)\beta[H(x) - U(\beta)] \\
&= -\frac{1}{Z(\beta)} \int_{\Gamma} dx \ [1 - (1-q)\beta[H(x) - U(\beta)]]_+^{1/(1-q)} + 1.
\end{aligned}
\tag{B1}
$$

Hence one obtains

$$
Z(\beta) = \int_{\Gamma} dx \ [1 - (1-q)\beta[H(x) - U(\beta)]]_+^{1/(1-q)} .
\tag{B2}
$$

7

From the latter expression follows

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}\beta} \ln Z(\beta) &= -\frac{1}{Z(\beta)} \int_\Gamma \mathrm{d}x \, [1 - (1-q)\beta[H(x) - U(\beta)]]_+^{q/(1-q)} \\
&\quad \times \left[H(x) - U(\beta) - \beta \frac{\mathrm{d}}{\mathrm{d}\beta} U(\beta) \right] \\
&= \beta \frac{\mathrm{d}}{\mathrm{d}\beta} U(\beta) \\
&= \frac{\mathrm{d}}{\mathrm{d}\beta} S(U(\beta)).
\end{aligned} \tag{B3}
$$

In the last line (7) has been used. Integration gives

$$
S(U(\beta)) = \ln Z(\beta) + \text{ constant.} \tag{B4}
$$

This is (8), up to the integration constant.

Let us now show that (8) in case of the ideal gas and in the limit $q = 1$ reduces to (9). For the ideal gas, in the limit $q = 1$, is

$$
\begin{aligned}
Z(\beta) &= e^{\beta U(\beta)} \int_\Gamma \mathrm{d}x \, e^{-\beta H(x)} \\
&= e^{\beta U(\beta)} V^n \left(\frac{2\pi}{\beta m} \right)^{\nu n/2}.
\end{aligned} \tag{B5}
$$

Therefore, (8) becomes

$$
S(U(\beta)) = \ln \left[\left(\frac{m}{h} \right)^{\nu n} e^{\beta U(\beta)} V^n \left(\frac{2\pi}{\beta m} \right)^{\nu n/2} \right]. \tag{B6}
$$

This yields (9) because $\beta U(\beta) = \nu n/2$.

References

[1] C. Tsallis, *Possible Generalization of Boltzmann-Gibbs Statistics*, J. Stat. Phys. **52**, 479-487 (1988).

[2] E.K. Lenzi, R.S. Mendes, L.R. da Silva, *Statistical mechanics based on Rényi entropy*, Physica A**280**, 337-345 (2000).

[3] J. Naudts, M. Czachor, *Thermostatistics based on Kolmogorov-Nagumo averages: Unifying framework for extensive and non-extensive generalizations*, Phys. Lett. A**298**, 369-374 (2002).

[4] A.G. Bashkirov, *On maximum entropy principle, superstatistics, power-law distribution and Renyi parameter,* to appear in Physica A (2004).

[5] P. Jizba, T. Arimitsu, *Generalized statistics: yet another generalization,* to appear in Physica A (2004).

[6] C. Tsallis, R.S. Mendes, A.R. Plastino, *The role of constraints within generalized nonextensive statistics,* Physica A**261**, 543-554 (1998).

[7] S. Martinez, F. Nicolás, F. Pennini, A. Plastino, *Tsallis' entropy maximization procedure revisited,* Physica A**286**, 489-502 (2000).

[8] H. B. Callen, *Thermodynamics and an introduction to thermostatistics* (John Wiley & Sons, 1985).

[9] S. Abe, *Thermodynamic limit of a classical gas in nonextensive statistical mechanics: negative specific heath and polytropism,* Phys. Lett. A**263**, 424-429 (1999); A**267**, 456-457 (2000).

[10] *E.T. Jaynes: Papers on probability, statistics and statistical physics.* ed. by R.D. Rosenkrantz (Kluwer Academic Publishers, 1989)

[11] E.M.F. Curado, C. Tsallis, *Generalized statistical mechanics: connection with thermodynamics,* J. Phys. A **24**, L69-72 (1991).

[12] C. Beck, F. Schlögl, *Thermodynamics of chaotic systems: An introduction* (Cambridge University Press, Cambridge, 1993)

[13] J. Naudts, *Generalized thermostatistics and mean field theory,* arXiv:cond-mat/0211444v6, Physica A**332**, 279-300 (2004).

[14] J. Naudts, *Generalized thermostatistics based on deformed exponential and logarithmic functions,* arXiv::cond-mat/0311438, to appear in Physica A.

[15] J. Naudts, *Estimators, escort probabilities, and phi-exponential families in statistical physics,* arXiv::math-ph/0402005.

[16] J. Naudts, *Rigorous results in non-extensive thermodynamics,* arXiv::math-ph/9908025, Rev. Math. Phys., **12**(10), 1305-1324 (2000).